Pattern Recognition Letters 32 (2011) 449–457

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Towards robust writer verification by correcting unnatural slant ☆

A.A. Brink [a,*], R.M.J. Niels [b], R.A. van Batenburg [a], C.E. van den Heuvel [c], L.R.B. Schomaker [a]

[a] Institute of Artificial Intelligence and Cognitive Engineering (ALICE), University of Groningen, P.O. Box 407, 9700 AK Groningen, The Netherlands
[b] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands
[c] Netherlands Forensic Institute, P.O. Box 24044, 2490 AA Den Haag, The Netherlands

ARTICLE INFO

Article history:
Received 11 September 2009
Available online 30 October 2010
Communicated by C.L. Tan

Keywords:
Handwriting biometrics
Writer verification
Slant
Disguise
Statistical pattern recognition

ABSTRACT

Slant is a salient feature of Western handwriting and it is considered to be an important writer-specific feature. In disguised handwriting however, slant is often modified. It was tested whether slant is indeed an important factor and it was tested whether the distorting effect of deliberate slant change can be countered by a simple shear transform. This was done in two off-line writer verification experiments in image processing conditions of slant elimination and slant correction. The experiments were performed using three features based on statistical pattern recognition, including the state-of-the-art features Fraglets and Hinge. A new public dataset was created and used, containing natural and slanted handwriting by 47 writers. A striking result is that the average natural slant value is much less important for biometric systems than is usually assumed: eliminating slant yields just a 1–5% performance loss. A second result is that the effects of deliberate slant change cannot be fully countered by a simple shear transform: it raises performance on the distorted handwriting from 53–68% to 64–90%, but this is still lower than normal operation on natural handwriting: 97–100%.

1. Introduction

A salient property of Western handwriting is *slant*: the dominant angle of near-straight downstrokes with respect to the horizontal. Slant is caused by the choice of pen grip and the relative contributions of wrist and finger movements. It has been modeled as the effect of locally using a single actuator (muscle) in a two-dimensional neuromuscular apparatus (Dooijes, 1986). Slant seems to be a key feature for writer verification: it plays an important role in biometric systems, as it is a major constituent of angular features (Bulacu and Schomaker, 2003; Crettez, 1995; Maarse, 1987). For example, the state-of-the-art Hinge feature (Bulacu and Schomaker, 2007) is based on angular frequencies; it is influenced by both curvature and slant. Furthermore, forensic document examiners and paleographers use this feature as a discriminatory characteristic (Burgers, 1995; Hardy and Fagel, 1995). These facts suggest that slant is a key feature for writer verification. However, it is not known to what extent slant contributes as an isolated factor to the performance of biometric systems for handwriting and its value may be overestimated.

In particular, slant is not a valuable feature in (possibly) disguised handwriting. In such a case, the handwriting was produced in a deliberately modified style, with the intention to avoid recognition of the writer's identity. Disguised handwriting is often used in threatening or stalking letters. In some cases, the mutilation of shapes successfully disturbs handwriting examination by forensic experts (Found and Rogers, 2005). Moreover, disguised handwriting cannot be handled by state-of-the-art systems for handwriting biometrics (writer verification and identification): computational features that are invariant to disguise do not exist. This is one of the reasons why systems for handwriting biometrics are not fully suitable for application in the forensic domain yet. Other unmet requirements are explainability of the system, robustness for variation in background effects, and robustness for forgery. Those issues have been addressed to some extent (Brink et al., 2007, 2008; Cha and Tappert, 2002; Franke and Köppen, 2000), but computational robustness against disguise is a largely untouched problem area.

A strategy to handle disguise is by applying an image operation to undo the effect of disguise, resulting in handwriting that is close to natural. This seems possible for the most frequently used kind of handwriting disguise: a change of slant. It is not surprising that slant modification is the most frequently used kind of disguise (Harris, 1953; Koppenhaver, 2007; Morris, 2000; Nickell, 2007), since humans can easily modify the slant during writing, and the effect on the visual appearance is dramatic (Koppenhaver, 2007). Therefore, an important step in making biometric systems robust for disguise is by correcting the slant. An obvious approach is to perform the correction by transforming the image with the *shear* operation, possibly resulting in the writer's natural handwriting.

☆ This research was made possible thanks to NWO Grant 634.000.434 (ToKeN/TriGraph).

* Corresponding author. Tel.: +31 (0) 55 576 7043; fax: +31 (0) 363 6687.
E-mail addresses: a.a.brink@ai.rug.nl (A.A. Brink), donders@ralphniels.nl (R.M.J. Niels), R.van.Batenburg@ai.rug.nl (R.A. van Batenburg), E.van.den.Heuvel@nfi.minjus.nl (C.E. van den Heuvel), schomaker@ai.rug.nl (L.R.B. Schomaker).

0167-8655/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2010.10.010

The objective of this study is twofold. The first objective is to determine how much information about the writer's identity is contained in the slant characteristic of natural handwriting. This will be tested in the first experiment by eliminating the slant in natural handwriting (*slant elimination*) and measuring to what extent the performance of automatic writer verification degrades. This experiment contributes to the theoretical basis of computational writer features based on directionality, such as the Hinge feature (Bulacu and Schomaker, 2007). The result will direct the design of future features.

The second objective is to determine the effectiveness of the shear transform in correcting handwriting disguised by slant change, when used as a preprocessing step before applying features such as Hinge (Bulacu and Schomaker, 2007) and Fraglets (Schomaker et al., 2004). Hinge and Fraglets are state-of-the-art features, based on statistical pattern recognition, which show impressive performance in test conditions.

At the same time, the underlying question will be answered: to what extent is a change of slant during human production of handwriting functionally equivalent to a shear transform? Slant change may result in more than just a shear effect, since it requires a non-habitual movement of the finger-wrist system, which may affect curvature. It has been suggested that there must also be an effect on writing speed, pressure, connecting strokes, style, construction, and size (Morris, 2000). Furthermore, disguised handwriting is less consistent (Harris, 1953; Koppenhaver, 2007; Morris, 2000). In the second experiment, it will be quantitatively determined to what extent such other effects occur. This will be done by shearing slanted text back to the supposed writer's natural slant angle (*slant correction*), and determining the performance of writer verification using state-of-the-art features. This is a first step in designing new biometric systems that are robust to disguise. To the best of our knowledge, no similar experiment has been performed before.

The experiments will be performed on a newly created public dataset: the *TriGraph slant dataset*, containing both natural and slanted handwriting of 47 subjects. It is described into more detail in the next section. In Sections 3 and 4, methods for slant estimation and feature extraction are described; these are preliminaries for the experiments. Experiment 1 will show that slant is not as informative as is usually assumed; it is described in Section 5. Experiment 2 will show that deliberate slant change is not equal to a simple shear transform; it is described in Section 6. Section 7 summarizes the conclusions.

2. TriGraph slant dataset

A new dataset was created, the *TriGraph slant dataset*: a unique collection of clean, deliberately slanted handwriting in conjunction with each writer's natural handwriting. It consists of 188 scanned images of handwritten pages, written by 47 untrained Dutch subjects, aged 27 on average. This dataset is relatively small compared to other datasets such as Firemaker (Schomaker and Vuurpijl, 2000) (251 writers), IAM (Marti and Bunke, 1999) (657) and Srihari's dataset (Srihari et al., 2002) (1500). However, the dataset proved to be large enough to analyze the effect of slant. It can be obtained from http://www.unipen.org/trigraphslant.html. The dataset can be used for both handwriting comparison experiments and handwriting recognition experiments.

The dataset was assembled as follows. The subjects were provided two printed Dutch texts, *text A* and *text B*. Both texts contained approximately 200 characters, including all lowercase letters and many capitals; the distribution of the letters among the two texts was similar. Each subject wrote four pages, such as the one shown in Fig. 1, following these instructions:

1. [*AN*] Copy text A in your natural handwriting.
2. [*BN*] Copy text B in your natural handwriting.
3. [*BL*] Copy text B and slant your handwriting to the *left* as much as possible.
4. [*BR*] Copy text B and slant your handwriting to the *right* as much as possible.

See Fig. 2 for a close look at fragments of the four pages written by one writer. The codes *AN*, *BN*, *BL* and *BR* refer to subsets into which the collected pages of the writers were subdivided. *AN* represents a collection of authentic documents; *BN*, *BL* and *BR* can be seen as collections of questioned documents. To avoid structural effects of fatigue, the order of item 3 and 4 was randomized; half of the subjects wrote the *BR* page before the *BL* page.

3. Slant estimation

Since Experiments 1 and 2 both require a reliable technique to estimate slant, a limited comparison of techniques is included here. A variety of slant estimation methods exists, based on different definitions of 'slant'. For example, it has been defined as the average direction of near-straight or long downstrokes (Maarse and Thomassen, 1983), or "the angle between the vertical direction and the direction of the strokes that, in an ideal model of handwriting, are supposed to be vertical" (Vinciarelli and Luettin, 2001). The methods can be roughly subdivided into two general approaches which could be called the *angle-frequency approach* (AF) and the *repeated-shearing approach* (RS). In AF, which is most popular (Kavallieratou et al., 2001), downstrokes are first located based on a criterion such as a minimal vertical extent or velocity. Next, the angle of the local ink direction is measured at those locations; the resulting angles are agglomerated in a histogram. From this histogram, the slant angle is determined. This general algorithm is shown in Algorithm 1. Variations include computing an edge-direction histogram and finding the maximum or mode in it (Bulacu and Schomaker, 2003) or the peak that is closest to 90° (Crettez, 1995). Another variation computes the average angle in rectangular sub-areas containing vertical structures (Bozinovic and Srihari, 1989).

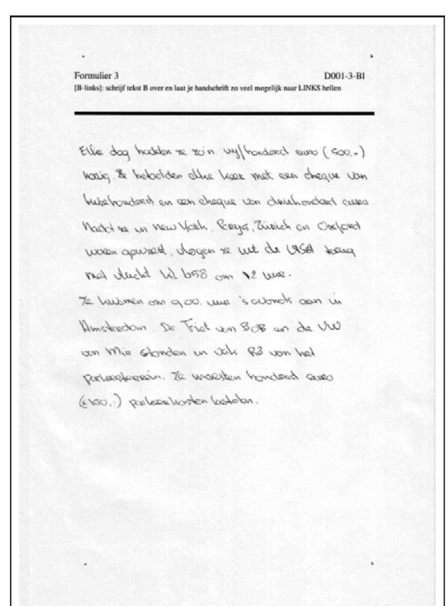
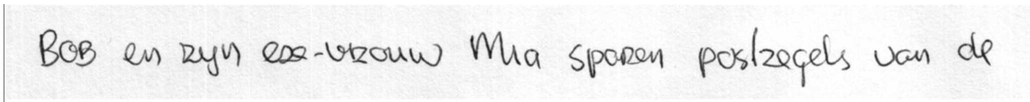
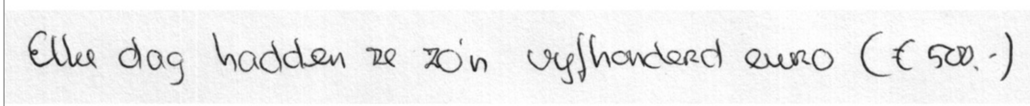


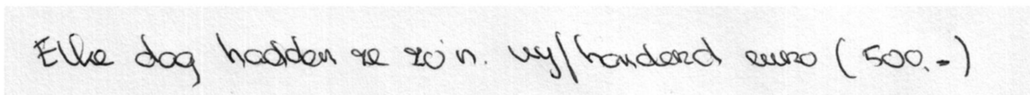
Fig. 1. Example page from the TriGraph slant dataset: page 3 of writer D001. It contains text B, slanted to the left (*BL*).



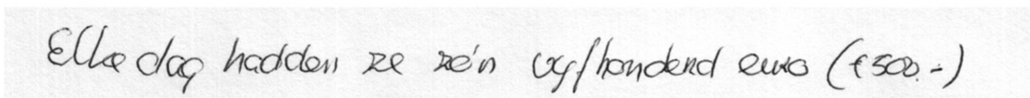
(a) Text A, natural handwriting (AN)



(b) Text B, natural handwriting (BN)



(c) Text B, slanted to the left (BL)



(d) Text B, slanted to the right (BR)

Fig. 2. Example of the four pages written by writer D001, the first subject in the dataset. For each page only the first line is shown, manually cut out for the purpose of illustration.

Algorithm 1. AF: Compute the slant angle using the angle-frequency approach. INPUT: image I. OUTPUT: slant angle a

```
h ← empty_histogram()
for all pixel p in I do
  if criterion(p) then
    a ← local_angle(p)
    h.add(a)
  end if
end for
return best_freq(h) {maximum or mode}
```

RS is based on the assumption that the projection of dark pixels is maximal along an axis parallel to the slant angle. The basic principle is to repeatedly shear images of individual text lines, varying the shear angle, and optimizing a criterion on the vertical projection of dark pixels (Kavallieratou et al., 2001; Vinciarelli and Luettin, 2001). Such a criterion involves the maximization of peaks in the projection. The range of shear angles extends to hypothetical extreme slant angles such as 30°···150°. This approach is shown in Algorithm 2. It requires that the text has been split into individual text lines, which can be done by using smoothed projection histograms if the text lines do not overlap much. Obviously, RS is much slower than AF, but that is of no importance for this experiment.

Algorithm 2. RS: Compute the slant angle using the repeated-shearing approach. INPUT: image I. OUTPUT: slant angle a

```
a ← empty_list()
for all textline L_i in I do
  s* = 0   for all a in 30°···150° do
    p ← ver_project_ink(shear(L,a))
    s ← score(p)
    if s > s* then
      s* ← s
      a_i ← a
    end if
  end for
end for
return median(a)
```

To determine a usable technique for slant estimation, a limited comparison of implementations of AF and RS was performed. AF was implemented by calculating the angle at near-straight parts of the ink boundary and yielding the mode of the smoothed histogram; RS was implemented with an algorithm that maximized the density of the 10 highest peaks in the vertical projection histogram of each text line. For testing purposes, ground truth data for the first 24 pages of the dataset was generated by averaging 10 manually measured downstroke angles per page; these were compared to the angle estimations of the automatic methods. Table 1 shows that the angles computed with RS are closer to the ground truth, although this difference is not significant at the 5% confidence level ($t = -1.12$; determined using the one-sample t-test: $t = \frac{\bar{x}}{s/\sqrt{n}}$, where $n = 24$, $\bar{x}$ is the average of x_i, which are the differences of squared errors, and s is the standard deviation of x). Still, in the following experiments, RS was used to automatically determine the slant angle.

4. Feature extraction and comparison

The effect of slant on features of handwriting was evaluated using three well-performing automatic features. These features

Table 1
Slant angle in the first 24 pages determined using three methods: manual (Man), angle frequencies (AF), and repeated shearing (RS). The manual measurements represent averages of at least 10 measurements per page. RS yielded the lowest RMSE (root of the mean of squared errors) with respect to the manually determined angle.

Page	Man	AF	RS	Page	Man	AF	RS
D001-1-AN	99	96	97	D004-1-AN	95	95	96
D001-2-BN	103	101	100	D004-2-BN	95	94	95
D001-3-BL	122	122	121	D004-3-BR	76	74	76
D001-4-BR	68	68	68	D004-4-BL	112	112	111
D002-1-AN	75	76	77	D005-1-AN	72	75	72
D002-2-BN	67	72	70	D005-2-BN	73	73	73
D002-3-BL	97	97	97	D005-3-BR	64	35	56
D002-4-BR	54	44	50	D005-4-BL	107	102	104
D003-1-AN	79	78	79	D006-1-AN	82	80	78
D003-2-BN	71	76	78	D006-2-BN	76	76	78
D003-3-BR	59	59	58	D006-3-BR	54	50	53
D003-4-BL	97	99	100	D006-4-BL	98	99	98
RMSE						7	3

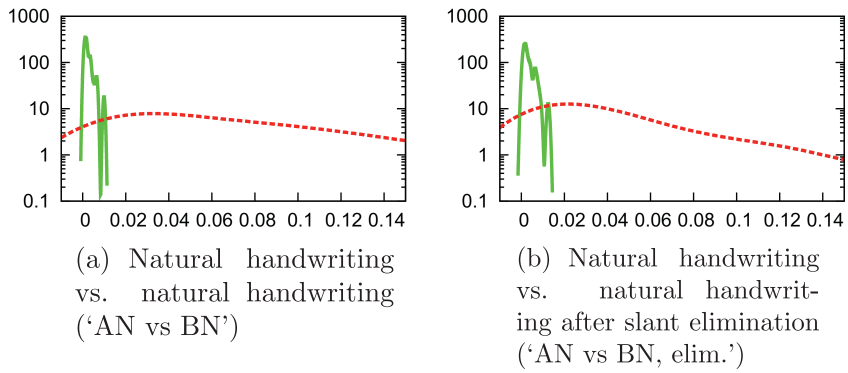


Fig. 3. Distribution of distances between documents with natural handwriting, based on the *Directions* feature. The continuous (green) curve represents D_s, distances between documents written by the same writer; the dashed (red) curve represents D_d, distances between documents by different writers. The two classes can be separated quite easily, either without (a) or with (b) slant elimination. For visualization purposes, the vertical axis is on log scale and the distributions were rendered smooth using Parzen windowing. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

will be briefly introduced below; refer to the respective papers for the details.

- The *Directions* feature (Bulacu and Schomaker, 2003) ($p(\phi)$) is a probability distribution (p.d.) of ink directions at the contours. This encodes slant and direction usage.
- The *Fraglets* feature (Schomaker et al., 2004) ($p(g)$, also named "fCO3") is a p.d. of usage of graphemes (fragments of handwriting) from a precomputed code book. This encodes allograph usage. *Fraglets* is one of the best features currently available.
- The *Hinge* feature (Bulacu and Schomaker, 2007) ($p(\phi_1, \phi_2)$) is a p.d. of angle combinations that are measured on the boundaries of the ink. This encodes slant and curvature. It is also one of the best features available.

The distance $d(\cdot,\cdot)$ between any two feature vectors $\mathbf{v}$ and $\mathbf{w}$ was computed with the χ^2 distance measure (Cha, 2007):

$$d(\mathbf{v}, \mathbf{w}) = \sum_{i=1}^{|\mathbf{v}|} \frac{(\mathbf{v}_i - \mathbf{w}_i)^2}{\mathbf{v}_i + \mathbf{w}_i} \tag{1}$$

where i is an index to the elements of $\mathbf{v}$ and $\mathbf{w}$. This distance measure emphasizes differences in small feature values. It was used for this experiment because it is specifically effective on feature vectors that represent a probability distribution (Schomaker and Bulacu, 2004), such as the three features described above. In the following, $d(P,Q)$ will denote the distance between the feature vectors of the images P and Q.

5. Experiment 1: information in slant

The first experiment focused on determining how informative the slant value in natural handwriting is. This was determined by computing the performance of writer verification on unmodified handwriting (denoted '*AN* vs *BN*'), and comparing it to the performance on handwriting of which the slant was eliminated ('*AN* vs *BN*, elim.'). This is explained in the next subsections.

5.1. Unmodified handwriting

For the performance on unmodified handwriting, documents were drawn from the dataset in pairs: one was drawn from the *AN* subset, the other from *BN*. After computing the distance between their feature vectors using one of the methods described in Section 4, two cases were distinguished: the documents were written by the same person, or by different persons. The same-writer distances formed a multiset of distances D_s and the different-writer distances formed the multiset D_d. Thus, D_s and D_d are defined as follows:

$$D_s = \{\forall i :: d(AN_i, BN_i)\}, \tag{2}$$
$$D_d = \{\forall i,j : i \neq j : d(AN_i, BN_j)\}. \tag{3}$$

Fig. 3 shows an example of D_s and D_d, visualized as distributions.

Based on these multisets of distances, a writer verification classifier was implemented by setting a threshold. The position of the threshold determines the trade-off between Type-I error rate (false accept rate; falsely assigning two pages to the same writer) and the Type-II error rate (false reject rate; falsely assigning two pages to different writers). It was put on the position where the Type-I error rate was equal to the Type-II error rate, or the equal-error rate (EER). The performance was estimated by $100\% \cdot (1 - \text{EER})$.

5.2. Slant-eliminated handwriting

To determine the contribution of slant to writer-specific features, the experiment was repeated after applying *slant elimination* on all pages: shearing the text such that its apparent slant becomes equal to 90°. It is a standard step in handwriting recognition systems, which use it to optimize recognition of the textual contents (Bertolami et al., 2007; Kavallieratou et al., 2001). It has also been called "deslanting", "slant removal", and "slant correction". In this way, the absolute slant information is lost. Slant elimination $E(\cdot)$ can be expressed as follows:

$$E(P) = \text{shear}(P, 90^\circ - a(P)), \tag{4}$$

where shear (P, α) is the image processing operation that shears a page image P with α degrees and $a(P)$ denotes the slant angle of the handwriting in P. $a(P)$ was estimated by the RS algorithm described in Section 3. Fig. 4 shows partial examples of slant-eliminated pages. Using slant elimination, the new definition of D_s and D_d is:

$$D_s = \{\forall i :: d(E(AN_i), E(BN_i))\}, \tag{5}$$
$$D_d = \{\forall i,j : i \neq j : d(E(AN_i), E(BN_j))\}. \tag{6}$$

The resulting performance $100\% \cdot (1 - EER)$ will be denoted '*AN* vs *BN*, elim.'

Notice that the pages were sheared entirely. An alternative option is to shear text lines or words individually, but this is less reliable and breaks ink traces at region boundaries. It is also possible to eliminate slant non-uniformly within each text line, but this seems to add little or no improvement, despite the added complexity (Bertolami et al., 2007). The page-level approach is simple, fast and keeps the signal structurally intact.

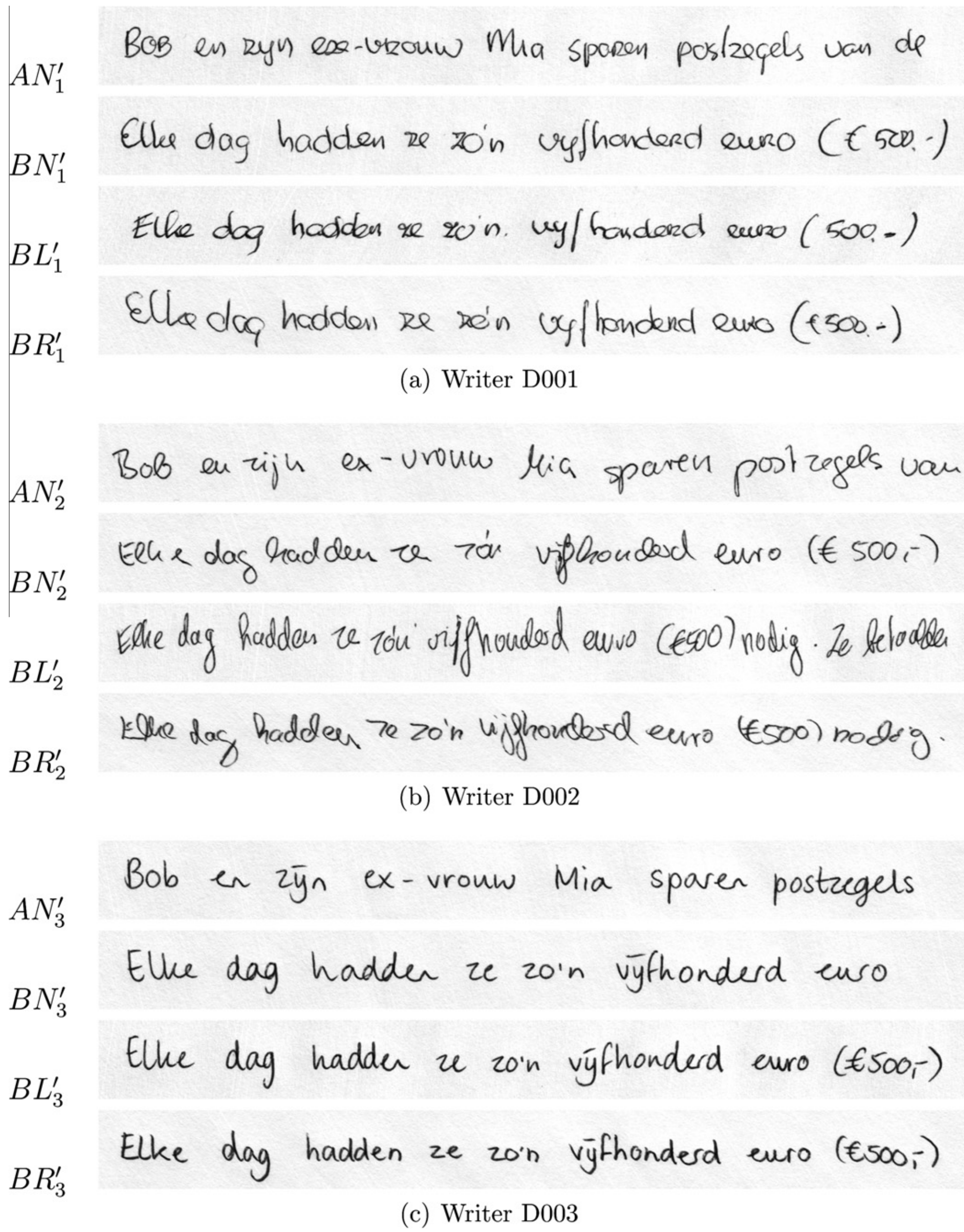


Fig. 4. Slant elimination in four pages by three writers; only the first text line of each page is shown. For each writer, four example text lines are shown: AN'_i, BN'_i, BL'_i and BR'_i. These are transformed images, obtained by executing slant elimination on the original pages. The originals written by subject D001 are shown in Fig. 2.

5.3. Results

The results of Experiment 1 are shown in Table 2. The first row shows performances of the three features on natural handwriting, '*AN* vs *BN*'. The performances of all features are 97% or higher, which confirms the high power of these features. These performances have an optimistic bias since they were not obtained on a separate test set, but the absolute performance is not relevant here. The second line of the table, where slant elimination was applied, shows a slight decrease of performance: for the Directions feature, it decreased with 5 percentage points. This is significant ($p \ll 0.001$, determined using the χ^2 test on the contingency table shown in Table 3) but small. The performance of the other features decreased with only 1 percentage point, a non-significant difference.

Table 2
Decrease of performance after slant elimination. Writer verification performance for three different features as the percentage of correct classifications. The value of slant seems very limited: only the Directions feature suffered somewhat, while the performances of the other features did not decrease significantly.

Subsets	Directions	Fraglets	Hinge
AN vs BN	97	100	99
AN vs BN, elim.	92	99	98

5.4. Discussion

Contrary to common assumptions, the results of Experiment 1 show that slant is not an important aspect of writer-specific features. Slant elimination did not significantly affect Fraglets and

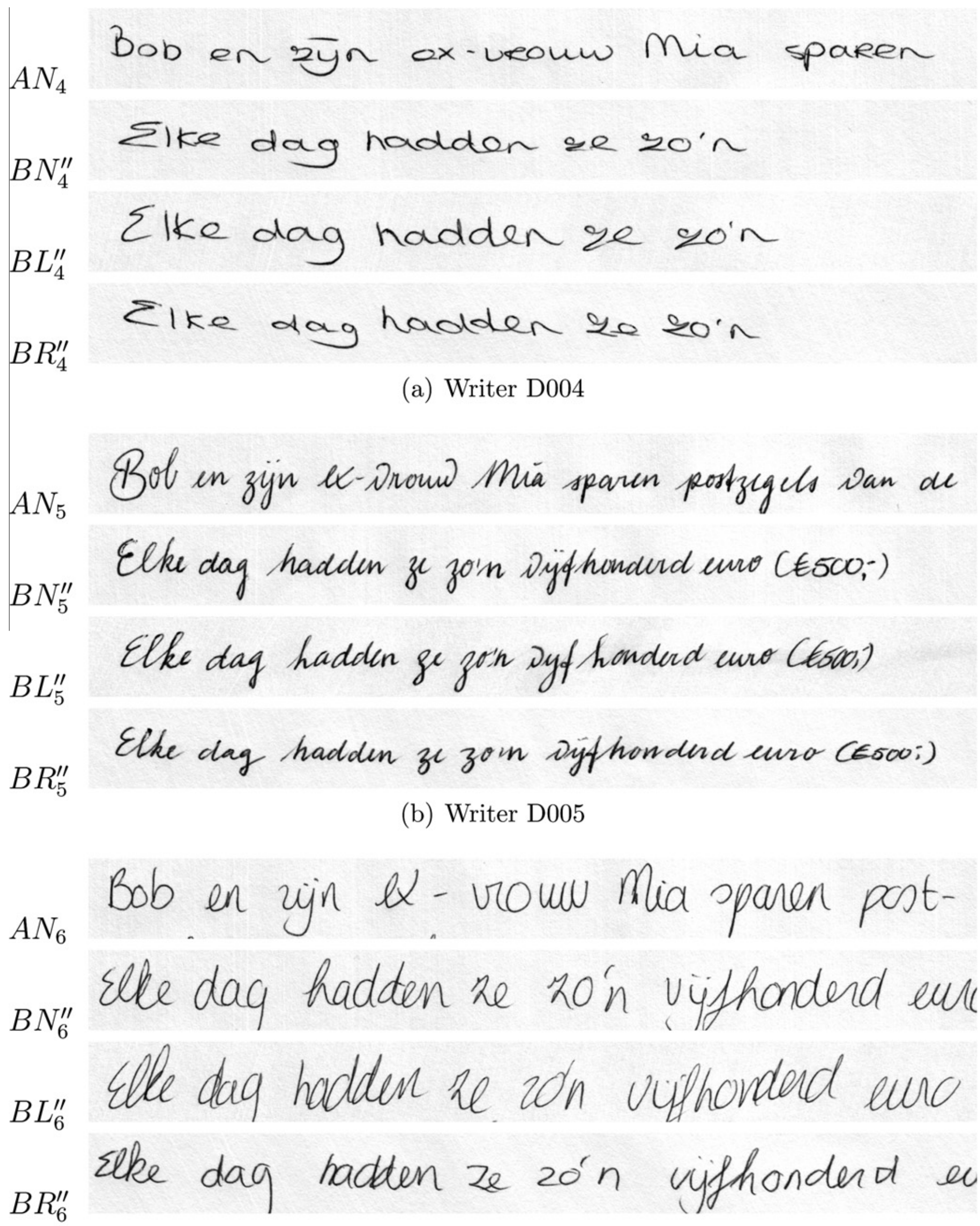
Table 3
Comparison of writer verification classifications using the *Directions* feature on natural handwriting in two conditions: with and without slant elimination. This contingency table elaborates on the found differences of performance shown in Table 2. The off-diagonal entries show the number of unequal classifications in the two conditions. In this case, the difference of the classifications in the two conditions is significant ($p \ll 0.001, \chi^2$ test of significance on this table); the Directions feature significantly performs slightly worse after slant elimination. The other features do not show a significant effect.

		Slant eliminated ('AN vs BN, elim.')	
		Correct	Wrong
Original	Correct	2015	145
('AN vs BN')	Wrong	6	43

Hinge, while the performance of Directions decreased with only 5%. The latter relies heavily on angular information; it is a distribution of angles in which the position of the mode (peak) indicates the average slant angle. The small decrease in its performance shows that the shape of the distribution is more important than its position. This also indicates that the shear transform can be used to counter slant changes in disguised handwriting. However, the next experiment shows that this is not completely effective.

6. Experiment 2: is deliberate slant change an affine transform?

The aim of the second experiment is to determine whether deliberate slant change is functionally equivalent to a simple affine transform: *shear*. In this experiment, apart from natural handwriting (*BN*), the disguised handwriting (*BL*, *BR*) from the dataset was included as well. Thus the experiment was performed three times, each time comparing documents from *AN* with those from either

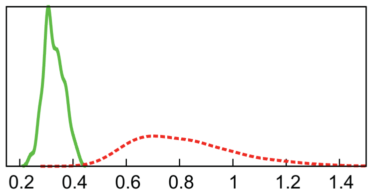


(a) Writer D004

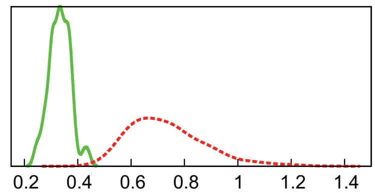
(b) Writer D005

(c) Writer D006

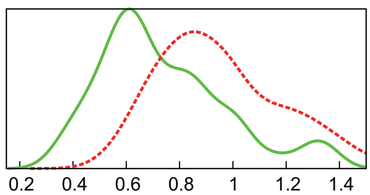
Fig. 5. Slant correction. For each of three writers, four example text lines are shown: AN_i, BN''_i, BL''_i and BR''_i. The latter three are transformed images, obtained by slant-correcting *BN*, *BL* and *BR*, respectively, with their slant matching that of the first line (AN).



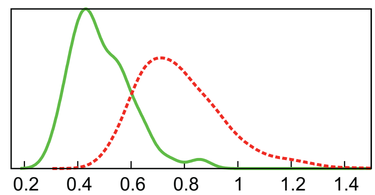
(a) Natural handwriting vs. natural handwriting ('AN vs BN')



(b) Natural handwriting vs. natural handwriting after correction ('AN vs BN, corr.'



(c) Natural handwriting vs. left-slanted handwriting ('AN vs BL')



(d) Natural handwriting vs. left-slanted handwriting after correction ('AN vs BL, corr.')

Fig. 6. Distribution of distances between documents with natural and slanted handwriting, based on the *Fraglets* feature. The continuous (green) curve represents D_s, distances between documents written by the same writer; the dashed (red) curve represents D_d, distances between documents by different writers. In the case of unmodified natural handwriting (a), the two classes can be separated easily. This changes when the writers disguise their handwriting by a slant change to the left (c). This is partly solved by slant correction (d), which does little harm to natural handwriting (b). For visualization purposes, the vertical axis is linear and the distributions were rendered smooth using Parzen windowing. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

BN, *BL* or *BR*. Furthermore, in an attempt to restore the handwriting, *slant correction* was used instead of slant elimination. This is explained into more detail in the next subsections.

6.1. Unmodified handwriting

The baseline performance for this experiment was computed similar to Experiment 1 (see Section 5.1). However, this time three baseline performances were computed: '*AN* vs *BN*', '*AN* vs *BL*' and '*AN* vs *BR*'.

6.2. Slant-corrected handwriting

In this experiment, the hypothesis is tested whether deliberate slant change is functionally a shear transform. If this is true then the manipulated handwriting can be transformed back to natural handwriting by shearing it such that the apparent slant becomes equal to the writer's natural slant angle. We define *slant correction* $C(\cdot,\cdot)$ as follows:

$$C(P,Q) = \text{shear}(P, a(Q) - a(P)). \quad (7)$$

It is quite similar to slant elimination, but this approach attempts to restore the original handwriting. It can be used if the handwriting in *Q* is known to be genuine but *P* may have been disguised by slant manipulation. This is illustrated in Fig. 5, which shows fragments of pages after slant correction.

In this condition, the distances were computed as follows:

$$D_s = \{\forall i :: d(AN_i, C(B_i, AN_i))\}, \quad (8)$$

$$D_d = \{\forall i,j : i \neq j : d(AN_i, C(B_j, AN_i))\}, \quad (9)$$

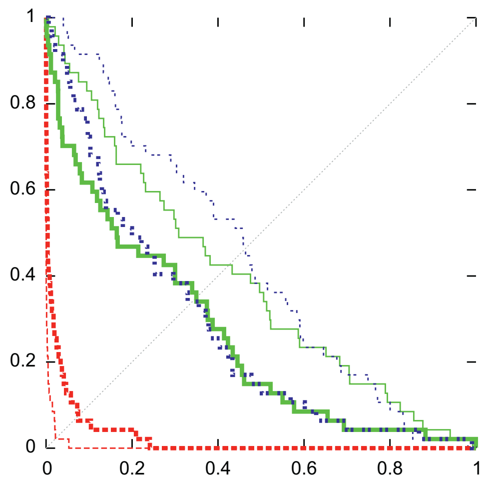
where *B* is either *BN*, *BL* or *BR*. Examples of the distributions of D_s and D_d are visualized in Fig. 6. The resulting performances are denoted '*AN* vs *BN*, corr.', '*AN* vs *BL*, corr.' and '*AN* vs *BR*, corr.' If the hypothesis is true, then these performances should be equal to the performance on natural handwriting ('*AN* vs *BN*').

6.3. Results

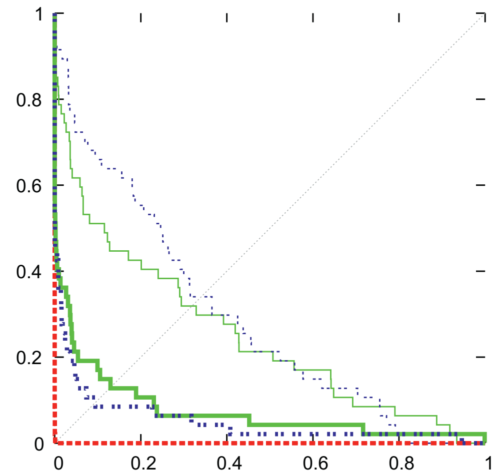
The results of Experiment 2 are shown in Table 4. The first row is a copy of the first row in Table 2. The best writer verification is obtained on natural handwriting ('*AN* vs *BN*'): for the tested features, performances are in the range 97–100%. The performances on natural vs slanted handwriting ('*AN* vs *BL*' and '*AN* vs *BR*') are obviously lower: a drop to 53–68%. These figures are all significantly differing from the corresponding performances on natural handwriting (χ^2 test, $p \ll 0.001$). "Correcting" the slant in *natural* text, which should not need correction ('*AN* vs *BN*, corr'), had only little negative influence on writer verification: Hinge and Fraglets remained stable, but the performance of the Directions feature dropped from 97% to 92% because it relies more on absolute slant information. But the most important result is that the performance on slant-corrected, slanted handwriting ('*AN* vs *BL*, corr' and '*AN* vs *BR*, corr') is significantly *lower* than the performance on natural handwriting ($p \ll 0.001$); the figures are only in the range 64–90%.

Table 4
Quantitative effect of slant manipulation on three writer-specific features. Writer verification performance on original images (first three lines) and slant-corrected images (last three lines). Percentages of correct classifications.

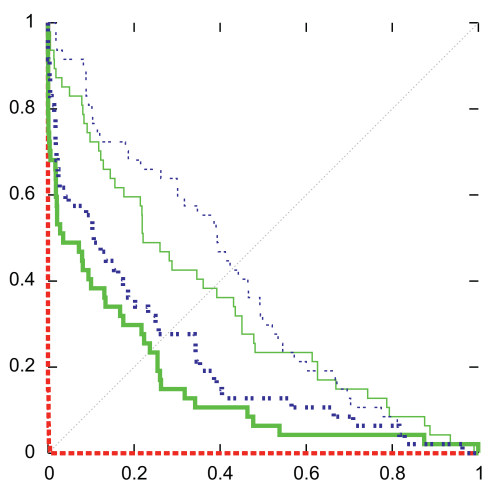
Subsets	Directions	Fraglets	Hinge
AN vs BN	97	100	99
AN vs BL	57	68	62
AN vs BR	53	66	57
AN vs BN, corr.	92	99	99
AN vs BL, corr.	64	86	75
AN vs BR, corr.	66	90	72



(a) Directions feature



(b) Fraglets feature



(c) Hinge feature

Fig. 7. Effectiveness of correcting deliberate slant change in handwriting; the key is in Fig. 8. Verification error plots for three features. The curves show the trade-off of the Type-I error rate on the horizontal axis and Type-II error on the vertical axis. The performances reported in Table 4 appear as the intersections of the curves with the diagonals. The more a curve approaches the lower left corner, the better the performance. Since Fraglets and Hinge perform 99–100% on natural handwriting ('*AN* vs *BN*') and on natural handwriting after correction ('*AN* vs *BN*, corr'), the corresponding curves are in the lower-left corner of the figure and cannot be discerned. The graphs show a consistent increase of performance after correcting slant, but it does not get as good as on natural handwriting.

The reported performances focus on the equal-error rate (EER), where the Type-I and Type-II error rates are equal. To explore the trade-off between the Type-I and Type-II error into more detail, Fig. 7 shows the errors as a result of varying the classification threshold. In the graphs, the EER values can be found at the intersection of any curve with the diagonal (shown as a dashed gray line. The key of Fig. 7 is in Fig. 8).

6.4. Discussion

The results of Experiment 2 show that handwriting disguise by changing slant lowers writer verification performance, if no correction is applied. This is obvious, but the effect is not the same on all tested features: the *Fraglets* feature seems to be most resilient against disguise by slant manipulation. After automatically correcting the slant with a shear operator, the performance improved for all tested features, which means that the distortion caused by slant change can be partly undone by slant correction. This result suggests to apply slant correction always before handwriting comparison takes place. The best performing feature after slant

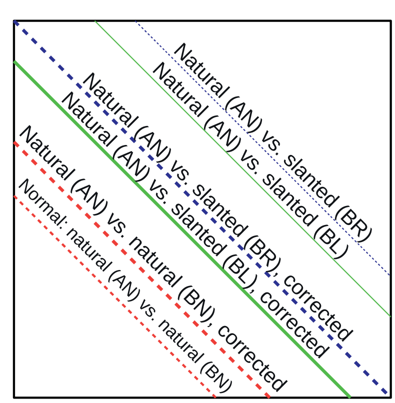


Fig. 8. Key for Fig. 7.

correction was Fraglets. But in spite of the improved performance, correcting the slant in slanted handwriting did not restore writer verification performance fully for any feature. This means that using the shear transform is not a complete solution against the problem of slant manipulation in disguised handwriting. In other words, slant correction did not result in the original handwriting, thus deliberate slant change is *not* functionally equivalent to the affine transform which is the shear operation.

This raises the question what else changes in the handwriting during slant change. We know that disguise is usually inconsistent (Harris, 1953; Koppenhaver, 2007; Morris, 2000), thus a greater variation of slant is expected. This is confirmed by the observation that Hinge and particularly Directions, which heavily rely on slant information, suffered most from the slant manipulation. Furthermore, a non-habitual movement of the finger-wrist motor system probably introduces artifacts. However, the dataset is not extensive enough for a conclusive analysis. Therefore, we suggest a follow-up study with more data, in which the corrected handwriting is thoroughly analyzed by forensic experts.

A further improvement would be to automatically *detect* disguise and decide if slant correction should be applied. Forensic experts try to detect disguise based on experience, but to the best of our knowledge, automatic methods to detect disguise do not exist yet. In one study, a method has been devised to approach the related problem of forgery detection (Cha and Tappert, 2002). It exploits the fact that forged handwriting is often less fluently written; this principle may be applicable to disguise detection as well. In addition, we suggest to exploit the aspect of inconsistency, as it is known to be an important indicator.

Another challenge for the future is to develop features that are invariant to disguise. A new direction might be to make features that determine the way the letters are constructed, mimicking an approach used by forensic document examiners.

7. Conclusion

Slant is a salient feature of handwriting and it is an important factor of state-of-the-art features, but as an isolated factor, it is not essential for good writer verification performance. It is not as informative for handwriting comparison as is usually assumed. This was found in a series of writer verification experiments using three state-of-the-art statistical features: Directions, Fraglets, and Hinge. Removing the absolute slant lowered writer verification performance by only 1–5 percentage points.

In disguised handwriting, slant is not valuable and possibly deceptive because it is subject to deliberate modification. When a non-habitual slant angle is applied during writing, performance of the features obviously decreased. However, correcting the handwriting by shearing it to obtain a natural slant value did not restore performance fully. Thus, disguise by slant change has more effect on the handwriting than just a shear effect, and the shear transform is not a complete solution against it. However, it is useful as a partial solution: slant correction did improve performance on disguised handwriting. Since slant is not an important aspect of handwriting, all possibly disguised handwriting can and should be sheared to match the specimen handwriting. This may be an essential step in biometric systems and manual comparisons as well.

References

Bertolami, R., Uchida, S., Zimmermann, M., Bunke, H., 2007. Non-uniform slant correction for handwritten text line recognition. In: Proc. 9th Internat. Conf. on Document Analysis and Recognition, pp. 18–22.

Bozinovic, R.M., Srihari, S.N., 1989. Off-line cursive script word recognition. IEEE Trans. Pattern Anal. Machine Intell. (PAMI) 11 (1), 68–83.

Brink, A., Schomaker, L., Bulacu, M., 2007. Towards explainable writer verification and identification using vantage writers. In: ICDAR, pp. 824–828.

Brink, A., van der Klauw, H., Schomaker, L., 2008. Automatic removal of crossed-out handwritten text and the effect on writer verification and identification. In: Yanikoglu, B., Berkner, K. (Eds.), Proc. Document Recognition and Retrieval XV, IS& T/SPIE Internat. Symp. on Electronic Imaging.

Bulacu, M., Schomaker, L., 2003. Writer style from oriented edge fragments. In: CAIP 2003. LNCS. Springer, pp. 460–469.

Bulacu, M., Schomaker, L., 2007. Text-independent writer identification and verification using textural and allographic features. IEEE Trans. Pattern Anal. Machine Intell. (PAMI) 29 (4), 701–717.

Burgers, J., 1995. De paleografie van de documentaire bronnen in Holland en Zeeland in de dertiende eeuw. Peeters, Leuven.

Cha, S., Tappert, C.C., 2002. Automatic detection of handwriting forgery. In: Proc. 8th IWFHR, pp. 264–267.

Cha, S.-H., 2007. Comprehensive survey on distance/similarity measures between probability density functions. Internat. J. Math. Models Methods Appl. Sci. 1, 300–307.

Crettez, J.-P., 1995. A set of handwriting families: Style recognition. In: Proc. 3rd ICDAR, pp. 489–494.

Dooijes, E., 1986. Decomposition of the pen displacement signal in the analysis of handwriting motorics. In: I.T. Young et al. (Ed.), Signal Processing III: Theories and Applications, pp. 1343–1345.

Found, B., Rogers, D., 2005. Problem types of questioned handwritten text for forensic document examiners. In: Advances in Graphonomics: Proc. 12th Biennial Conf. Internat. Graphonomics Society, pp. 8–12.

Franke, K., Köppen, M., September 2000. A framework for document pre-processing in forensic handwriting analysis. In: Schomaker, L., Vuurpijl, L. (Eds.), Proc. 7th Internat. Workshop Frontiers Handwriting Recognition, pp. 73–82. <citeseer.ist.psu.edu/franke00framework.html>.

Hardy, H., Fagel, W., 1995. Methodological aspects of handwriting identification. J. Forensic Document Exam. 8, 33–69.

Harris, J.J., 1953. Disguised handwriting. J. Criminal Law, Criminol. Police Sci. 43, 685–689.

Kavallieratou, E., Fakotakis, N., Kokkinakis, G., 2001. Slant estimation algorithm for OCR systems. Pattern Recognition 34, 2515–2522.

Koppenhaver, K.M., 2007. Forensic Document Examination: Principles and Practice. Humana Press. <www.springerlink.com/content/r548g18095452072/>.

Maarse, F., 1987. The Study of Handwriting Movement: Peripheral Models and Signal Processing Techniques. Ph.D. Thesis, Katholieke Universiteit Nijmegen.

Maarse, F.J., Thomassen, A.J., 1983. Produced and perceived writing slant: Difference between up and down strokes. Acta Psychol. 54, 131–147.

Marti, U., Bunke, H., 1999. A full english sentence database for off-line handwriting recognition. In: Proc. 5th ICDAR, pp. 705–708.

Morris, R.N., 2000. Forensic Handwriting Identification. Academic Press.

Nickell, J., 2007. Detecting Forgery: Forensic Investigation of Documents. University Press of Kentucky.

Schomaker, L., Bulacu, M., 2004. Automatic writer identification using connected-component contours and edge-based features of uppercase western script. IEEE Trans. Pattern Anal. Machine Intell. 26 (6), 787–798.

Schomaker, L., Bulacu, M., Franke, K., 2004. Automatic writer identification using fragmented connected-component contours. In: Kimura, F., Fujisawa, H. (Eds.), 9th IWFHR. Tokyo, Japan, pp. 185–190.

Schomaker, L., Vuurpijl, L., 2000. Forensic writer identification: A benchmark data set and a comparison of two systems. Tech. rep., NICI, Nijmegen.

Srihari, S.N., Cha, S.-H., Arora, H., Lee, S., 2002. Individuality of handwriting. J. Forensic Sci. 47, 856–872.

Vinciarelli, A., Luettin, J., 2001. A new normalization technique for cursive handwritten words. Pattern Recognition Lett. 22, 1043–1050.